CAMBRIDGE

Cambridge University Press
978-1-107-00731-4 - Convexity: An Analytic Viewpoint
Barry Simon
Excerpt
More information

1

Convex functions and sets

This chapter has the fundamental definitions and some of the basics concerning differentiability and Jensen's inequality that will play central roles throughout the book. We'll also define the gauge of a convex set and Legendre transforms of functions, two notions central to Chapters 2–5. And we'll phrase the Hahn–Banach theorem as essentially a statement about tangents to convex functions.

A function, f, from an interval, $I \subset \mathbb{R}$, to $\mathbb{R}$ is called *convex* if and only if for all $x, y \in I$ and $\theta \in [0, 1]$,

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y) \tag{1.1}$$

Geometrically (see Figure 1.1), (1.1) says for z in the interval $[x, y]$, the pairs $(z, f(z))$ lie below the straight line from $(x, f(x))$ to $(y, f(y))$. It is remarkable that such a simple definition is so useful and rich. In particular, we will see that both Hölder's and Minkowski's inequalities are consequences of convex machinery – so much so that we will provide four distinct proofs of Hölder's inequality, three in this chapter and one in the next.

An equivalent formula to (1.1) is that for $x, y, z \in I$ with $x < y < z$, we have the determinant

$$\begin{vmatrix} x & f(x) & 1 \\ y & f(y) & 1 \\ z & f(z) & 1 \end{vmatrix} \geq 0$$

This is a geometric statement about a triangle being positively oriented.

To extend the definition from $\mathbb{R}$ to $\mathbb{R}^\nu$ (and beyond), we need to begin with domains of definition that generalize the role of intervals.

Definition A subset K of a real vector space, V, is called *convex* if and only if for any $x, y \in K$ and $\theta \in [0, 1]$, $\theta x + (1-\theta)y \in K$.

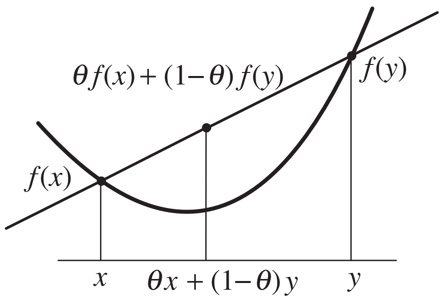

Figure 1.1 The meaning of a convex function

Thus, K is convex if it contains the line segment between any pair of points in K.

Definition Let K be a convex subset of a vector space V. A function $f\colon K \to \mathbb{R}$ is called

(i) *convex* if (1.1) holds for all $x, y \in K$ and $\theta \in [0,1]$,
(ii) *concave* if $-f$ is convex,
(iii) *affine* if f is convex and concave,
(iv) *strictly convex* if f is convex and strict inequality holds in (1.1) whenever $x \neq y$ and $\theta \in (0,1)$.

Thus, concavity means that

$$f(\theta x + (1-\theta)y) \geq \theta f(x) + (1-\theta) f(y)$$

and affine means

$$f(\theta x + (1-\theta)y) = \theta f(x) + (1-\theta) f(y)$$

If $K = V$ and f is affine with $f(0) = 0$, then first

$$f(\theta x) = f(\theta x + (1-\theta)0) = \theta f(x)$$

and then

$$f(x+y) = 2f(\tfrac{1}{2}\,x + \tfrac{1}{2}\,y) = f(x) + f(y)$$

so f is linear. In general, if $K = V$, every affine function f is of the form $f(x) = f(0) + \ell(x)$ with ℓ linear.

For some purposes in later chapters, it is convenient to allow a convex function to take the value $+\infty$ and to extend f from $K \subset V$ to all of V by setting it to ∞ on $V \backslash K$. Since K is convex, (1.1) then still holds for all $x, y \in V$. In this chapter though, we will suppose $f < \infty$ at all points of definition.

The following connection between convex sets and functions is easy to check:

Proposition 1.1 *Let K be a convex subset of V and $f: K \to \mathbb{R}$. Define*

$$\tilde{\Gamma}(f) = \{(x, \lambda) \in V \times \mathbb{R} \mid x \in K,\ \lambda > f(x)\} \tag{1.2}$$

Then f is a convex function if and only if $\tilde{\Gamma}(f)$ is a convex subset of $V \times \mathbb{R}$.

Moreover, a simple induction together with

$$\sum_{j=1}^{m} \theta_j x_j = \theta_m x_m + (1 - \theta_m) \sum_{j=1}^{m-1} \varphi_j x_j$$

where $\varphi_j = \theta_j (1 - \theta_m)^{-1}$ shows that

Proposition 1.2 (First Form of Jensen's Inequality) *If K is a convex subset of V and $x, \dots, x_m \in K$ and $\theta_1, \dots, \theta_m \in [0, 1]$ with $\sum_{j=1}^m \theta_j = 1$, then $\sum_{j=1}^m \theta_j x_j \in K$. If $f: K \to \mathbb{R}$ is convex, then*

$$f\left(\sum_{j=1}^{m} \theta_j x_j\right) \leq \sum_{j=1}^{m} \theta_j f(x_j) \tag{1.3}$$

The following is sometimes useful:

Proposition 1.3 *Let $f: I \subset \mathbb{R} \to \mathbb{R}$ be continuous. Then f is convex if and only if for all $x, y \in I$,*

$$f(\tfrac{1}{2}\, x + \tfrac{1}{2}\, y) \leq \tfrac{1}{2}\, f(x) + \tfrac{1}{2}\, f(y) \tag{1.4}$$

Remarks 1. (1.4) is called *midpoint convexity*.

2. Since convexity is a statement about f restricted to straight lines, the result immediately extends to f defined on $K \subset \mathbb{R}^\nu$ and on $K \subset V$, any vector space with a topology in which scalar multiplication and addition are continuous functions.

Proof Obviously, convexity implies (1.4). Suppose conversely that (1.4) holds. Write $\frac{1}{4}x + \frac{3}{4}y = \frac{1}{2}(\frac{1}{2}x + \frac{1}{2}y) + \frac{1}{2}y$ and conclude that

$$\begin{aligned} f(\tfrac{1}{4}\, x + \tfrac{3}{4}\, y) &\leq \tfrac{1}{2}\, f(\tfrac{1}{2}\, x + \tfrac{1}{2}\, y) + \tfrac{1}{2}\, f(y) \\ &\leq \tfrac{1}{4}\, f(x) + \tfrac{3}{4}\, f(y) \end{aligned}$$

By a simple induction, (1.1) holds for all dyadic rationals $\theta = j/2^n$, $j = 0, 1, 2, \dots, 2^n$. Then, by continuity, it holds for all θ. □

The following more complicated result will be convenient in Chapter 13:

Proposition 1.4 *Let $f: I \subset \mathbb{R} \to \mathbb{R}$ be lsc. Then f is convex if and only if for all $x, y \in I$,* (1.4) *holds.*

Remark lsc is short for lower semicontinuous, that is, if $x_n \to x$, then $f(x) \leq \liminf f(x_n)$.

Proof Let $[a,b] \subset I$ be a bounded interval. Since f is lsc, it is bounded below and takes its minimum value at a point $c \in [a,b]$ (for let $\alpha = \inf_{x\in[a,b]} f(x)$, let x_n be a sequence with $f(x_n) \to \alpha$, and let c be a limit point of the x_n). We will prove continuity on $[c,b]$. The proof for $[a,c]$ is similar, once one has that Proposition 1.3 applies. For notational simplicity, suppose $c = 0$ and $b = 1$.

By (1.4) and $f(0) \le f(1)$, $f(\frac{1}{2}) \le f(1)$, and since $f(0)$ is the minimum, $f(0) \le f(\frac{1}{2})$. Once one has this, $f(0) \le f(\frac{1}{4}) \le f(\frac{1}{2})$ and then since $f(\frac{1}{2}) \le \frac{1}{2}[f(\frac{1}{4}) + f(\frac{3}{4})]$, we conclude $f(\frac{1}{2}) \le f(\frac{3}{4})$, and then by (1.4), $f(\frac{3}{4}) \le f(1)$. By induction for any pair of x, y in $\mathbb{D}$, the dyadic rationals in $[0,1]$, $x < y \Rightarrow f(x) \le f(y)$.

It follows for any $y \in [0,1]$,

$$\tilde{f}(y) = \lim_{\substack{x\in\mathbb{D} \\ x\uparrow y}} f(x) = \sup_{\substack{x\in\mathbb{D} \\ x\uparrow y}} f(x)$$

exists. Notice $\tilde{f}$ is monotone and

$$\tilde{f}(y) = \lim_{x\uparrow y} \tilde{f}(x) \tag{1.5}$$

We claim $\tilde{f}(y)$ is continuous, for pick $x_n \uparrow y$ and $z_n \downarrow y$ with $x_n, z_n \in \mathbb{D}$ and arrange that $\frac{1}{2}(x_n + z_n) \ge y$. Then, with $\tilde{f}(y+) \equiv \lim_{x\downarrow y} \tilde{f}(x) = \lim_{x\downarrow y,\, x\in\mathbb{D}} f(x)$, we have by (1.4) that

$$\begin{aligned} \tilde{f}(y+) &= \lim_{n\to\infty} \tilde{f}(\tfrac{1}{2}\, x_n + \tfrac{1}{2}\, z_n) \\ &\le \lim_{n\to\infty} \tfrac{1}{2}\, \tilde{f}(x_n) + \tfrac{1}{2}\, \tilde{f}(z_n) \\ &= \tfrac{1}{2}\, \tilde{f}(y) + \tfrac{1}{2}\, \tilde{f}(y+) \end{aligned}$$

so $\tilde{f}(y+) \le \tilde{f}(y)$ which means, by monotonicity and (1.5), that $\tilde{f}$ is continuous.

By the lsc hypothesis,

$$f(y) \le \tilde{f}(y) \tag{1.6}$$

On the other hand, pick $x_n \uparrow y$ with $x_n \in \mathbb{D}$ and let $z_n = y + 2(x_n - y)$ so $x_n = \frac{1}{2} z_n + \frac{1}{2} y$ and by (1.4),

$$f(x_n) \le \tfrac{1}{2}\, f(z_n) + \tfrac{1}{2}\, f(y) \le \tfrac{1}{2}\, \tilde{f}(z_n) + \tfrac{1}{2}\, f(y)$$

by (1.6). Taking $n \to \infty$ and using the continuity of $\tilde{f}$, we see that

$$\tilde{f}(y) \le f(y)$$

Thus, $f = \tilde{f}$ is continuous and so convex. □

Remark In the proof of Proposition 9.15, we will see another situation where midpoint convexity implies convexity, namely, if f is monotone.

In the following, the use of Proposition 1.3 is of purely notational simplicity; one can directly deal with general θ.

Theorem 1.5 *Let $f\colon I \to \mathbb{R}$ with I an open interval and let f be C^2. Then f is convex if and only if*

$$f''(x) \geq 0 \tag{1.7}$$

for all $x \in I$. If K is an open convex subset of $\mathbb{R}^\nu$ and f is C^2 on K, then f is convex if and only if the Hessian $\partial^2 f/\partial x^i \partial x^j$ is positive definite at each point.

Remark This result is extended to f's which are not C^2 in Theorem 1.29.

Proof Consider first the case $\nu = 1$. Taylor's theorem with remainder says that for $\delta x > 0$,

$$f(x \pm \delta x) = f(x) \pm \delta x f'(x) + \int_0^{\delta x} (\delta x - y) f''(x \pm y)\, dy \tag{1.8}$$

and thus,

$$\tfrac{1}{2}\,[f(x+\delta x)+f(x-\delta x)]-f(x) = \tfrac{1}{2}\int_0^{\delta x} (\delta x-y)[f''(x+y)+f''(x-y)]\, dy \tag{1.9}$$

It follows that if (1.7) holds, then f obeys (1.4), and so f is convex by Proposition 1.3. Conversely, if f is convex, the left side of (1.9) is nonnegative for each x and each sufficiently small δx. Thus, taking the right side of (1.9) and dividing by $\frac{1}{2}(\delta x)^2$ and taking $\delta x \downarrow 0$, we see that (1.7) holds. We have thus proven the result if $\nu = 1$.

For general ν, we note that convexity is a statement about the values of f restricted to line segments in K. Thus, f is convex on K if and only if for all $x_0 \in K$ and $e \in \mathbb{R}^\nu$, $e \neq 0$, if we define $I_e(x_0) = \{\lambda \in \mathbb{R} \mid x_0 + \lambda e \in K\}$ and $F(\lambda; x_0, e) = f(x_0 + \lambda e)$, then F is convex as a function on $I_e(x_0)$. From the one-dimensional case, we see that F is convex if and only if $F''(\lambda) \geq 0$ for such λ. Since $F(\lambda; x_0, e) = F(\lambda - \lambda_0; x_0 + \lambda_0 e, e)$, we see f is convex if and only if for each x_0 and $e \neq 0$,

$$F''(0; x_0, e) \geq 0 \tag{1.10}$$

Since

$$F''(0; x_0, e) = \sum_{i,j=1}^{\nu} e_i e_j \,\frac{\partial^2 f}{\partial x_i \partial x_j}\,(x_0)$$

(1.10) is equivalent to the positive definiteness of the Hessian, as claimed. □

Remark The proof shows if $f''(x) > 0$ (indeed, if f'' is a.e. strictly positive), then f is strictly convex. The example $f(x) = x^4$, which is strictly convex but has $f''(0) = 0$, shows the converse is not true as a pointwise statement.

Cambridge University Press
978-1-107-00731-4 - Convexity: An Analytic Viewpoint
Barry Simon
Excerpt
More information

Example 1.6 Let $f(x) = e^x$ on $(-\infty, \infty)$. Then $f''(x) = e^x > 0$ so f is convex. Midpoint convexity

$$e^{\frac{1}{2}(x+y)} \leq \tfrac{1}{2}\, e^x + \tfrac{1}{2}\, e^y$$

is (if $a = e^x$, $b = e^y$ so a, b are arbitrary numbers on $(0, \infty)$)

$$\sqrt{ab} \leq \tfrac{1}{2}\,(a + b) \tag{1.11}$$

the arithmetic-geometric mean inequality. Thus, convexity of e^x generalizes this inequality. Since midpoint convexity implies convexity, (1.11) actually implies convexity of $x \to e^x$.

Using Proposition 1.2 with $\theta_1 = \cdots = \theta_m = 1/m$ and this f, we see that if $a_1, \dots, a_m > 0$, then

$$(a_1 \dots a_m)^{1/m} \leq \frac{1}{m}\,(a_1 + \cdots + a_m) \tag{1.12}$$

The function $g(x) = \log x$ for $x \in (0, \infty)$ obeys $g''(x) = -1/x^2 < 0$ so g is concave. By the above remark, f is strictly convex and g is strictly concave.

It is no coincidence that $\log x$, the inverse of e^x is concave. If f is strictly monotone and f is convex, then f^{-1}, the inverse function, is concave. For let x, y be given in $\operatorname{Ran} f$ and let a, b be chosen so that $x = f(a)$, $y = f(b)$. Then, convexity of f implies that

$$\theta x + (1-\theta)y \geq f((1-\theta)a + \theta b) \tag{1.13}$$

Since f is monotone, so is f^{-1}, so applying f^{-1} to (1.13), inequalities are preserved and thus,

$$\begin{aligned} f^{-1}(\theta x + (1-\theta)y) &\geq (1-\theta)a + \theta b \\ &= (1-\theta)f^{-1}(x) + \theta f^{-1}(y) \end{aligned}$$

that is, f^{-1} is concave. □

Proposition 1.7 *Let $f\colon [0, a] \to \mathbb{R}$ be convex and monotone increasing. Then $g\colon \{x \in \mathbb{R}^\nu \mid |x| \leq a\} \to \mathbb{R}$ by*

$$g(x) = f(|x|)$$

is a convex function.

Proof

$$\begin{aligned} g(\theta x + (1-\theta)y) &= f(|\theta x + (1-\theta y)|) \\ &\leq f(\theta|x| + (1-\theta)|y|) \\ &\leq \theta g(x) + (1-\theta)g(y) \end{aligned} \tag{1.14}$$

(1.14) follows from the assumed monotonicity of f and the triangle inequality

$$|\theta x + (1-\theta)y| \leq \theta|x| + (1-\theta)|y| \qquad \square$$

Remarks 1. The proof shows that $x \mapsto f(\|x\|)$ is convex on the ball of radius a in any normed linear space.

2. The proof also shows if f on $\mathbb{R}$ is even and convex, then f is monotone increasing on $[0,\infty)$.

Example 1.8 Let $p \geq 1$ and let $f(x) = |x|^p$ on $\mathbb{R}$. By the last proposition, for f to be convex, we only need that f is convex on $[0,\infty)$. By continuity, convexity on $(0,\infty)$ suffices and for that, we need only note that $f''(x) = p(p-1)x^{p-2} \geq 0$ so f is convex if $p \geq 1$ and strictly convex if $p > 1$. $\square$

Convexity and the triangle inequality are intimately related:

Theorem 1.9 *Let V be a vector space and let $F\colon V \to [0,\infty)$ be homogeneous of degree* 1*, that is,*

$$F(\lambda x) = \lambda F(x) \tag{1.15}$$

for all $x \in V$ and $\lambda \in [0,\infty)$. Then the following are equivalent:

(i) *F is convex.*
(ii) *$\{x \mid F(x) \leq 1\}$ is a convex set.*
(iii)

$$F(x+y) \leq F(x) + F(y) \tag{1.16}$$

In particular, F obeying (1.15) *is a seminorm if and only if F is convex with*

$$F(-x) = F(x) \tag{1.17}$$

and a norm if and only if F is convex, strictly positive on $V\backslash\{0\}$, and (1.17) *holds.*

Proof (i) ⇒ (ii) If F is convex and $x, y \in K \equiv \{z \mid F(z) \leq 1\}$, then $F(\theta x + (1-\theta)y) \leq \theta F(x) + (1-\theta)F(y) \leq 1$ so $\theta x + (1-\theta)y \in K$, that is, K is convex.

(ii) ⇒ (iii) Consider first the case $F(x) \neq 0 \neq F(y)$. Let $\theta = F(x)/[F(x) + F(y)]$. Since $x/F(x), y/F(y) \in K \equiv \{z \mid F(z) \leq 1\}$, (ii) implies $\theta x + (1-\theta)y = (x+y)/[F(x)+F(y)]$ lies in K, that is, $F(x+y)/[F(x)+F(y)] \leq 1$, that is, (1.16) holds.

If $F(x) = 0$ and $F(y) = 1$, then for any $\lambda > 0$, $\lambda x, y \in K$ so $\alpha_\lambda(x+y) = \theta_\lambda(\lambda x) + (1-\theta_\lambda)y$ with $\theta_\lambda = 1/(1+\lambda)$ and $\alpha_\lambda = \lambda/(1+\lambda)$. Thus,

$$\frac{\lambda}{1+\lambda} F(x+y) \leq 1$$

so taking $\lambda \to \infty$, $F(x+y) \leq 1$, so (1.16) holds. If $F(x) = 0$ and $F(y) \neq 0$, repeat the argument with x replaced by $x/F(y)$ and y by $y/F(y)$.

If $F(x) = F(y) = 0$, then for any $\lambda > 0$, $\lambda x, \lambda y \in K$ so $\frac{1}{2}\lambda(x+y) \in K$ so $F(x+y) \le 2/\lambda$. Taking $\lambda \to \infty$, $F(x+y) = 0$, so (1.16) again holds.

(iii) ⇒ (i) If (iii) holds, then

$$\begin{aligned} F(\theta x + (1-\theta)y) &\le F(\theta x) + F((1-\theta)y) && \text{(by (1.16))} \\ &= \theta F(x) + (1-\theta)F(y) && \text{(by (1.15))} \end{aligned}$$ □

Corollary 1.10 *Let K be a convex subset of V which obeys*
(i) *For any $x \in V$, $\lambda x \in K$ for some $\lambda > 0$.*
(ii) *If $x \in K$, then $-x \in K$.*
Define

$$\|x\| = \inf\{\lambda \in (0,\infty) \mid \lambda^{-1}x \in K\} \tag{1.18}$$

Then $\|\cdot\|$ is a seminorm on V. Moreover,

$$\{x \mid \|x\| \le 1\} = \bigcap_{\lambda > 1} \lambda K \tag{1.19}$$

$$\{x \mid \|x\| < 1\} = \bigcup_{\lambda < 1} \lambda K \tag{1.20}$$

Remarks 1. When (i) holds, we say that K is *absorbing*. When (ii) holds, we say K is *balanced*.

2. Thus, $\{x \mid \|x\| < 1\} \subset K \subset \{x \mid \|x\| \le 1\}$. But see Remark 5 below.

3. For any convex set K with $0 \in K$, the function of the right side of (1.18) is called the *gauge* of K.

4. By (ii), $0 = \frac{1}{2}(x - x) \in K$, so $\{\mu \mid \mu x \in K\}$ is a symmetric interval I. $\|x\|$ is defined so $\sup_{\mu \in I} \mu = \|x\|^{-1}$.

5. If V is $\mathbb{R}^\nu$ and K is open (resp. closed), then $K = \{x \mid \|x\| < 1\}$ (resp. $\{x \mid \|x\| \le 1\}$). More generally, if for all x, $\{\lambda \mid \lambda x \in K\} \subset \mathbb{R}$ is open, $K = \{x \mid \|x\| < 1\}$, and if the set is closed, then $K = \{x \mid \|x\| \le 1\}$.

6. If V is a complex vector space and (ii) is replaced by $x \in K$ and $|\zeta| = 1$ (for $\zeta \in \mathbb{C}$), then $\zeta x \in K$, then $\|\cdot\|$ is a complex seminorm.

Proof By (i), $\{\lambda \mid \lambda^{-1}x \in K\}$ is nonempty so $\|\cdot\|$ is everywhere defined. Clearly, if $\mu > 0$,

$$\begin{aligned} \|\mu x\| &= \inf\{\lambda \mid \lambda^{-1}\mu x \in K\} \\ &= \inf\{\lambda\mu \mid (\lambda\mu)^{-1}\mu x \in K\} \\ &= \inf\{\lambda\mu \mid \lambda^{-1}x \in K\} = \mu\|x\| \end{aligned}$$

so $\|\cdot\|$ is homogeneous of degree 1. Moreover, by (ii), $\|-x\| = \|x\|$.

Now

$$\begin{aligned}\{x \mid \|x\| \le 1\} &= \{x \mid \inf\{\lambda \mid \lambda^{-1}x \in K\} \le 1\} \\ &= \{x \mid \lambda^{-1}x \in K \text{ for all } \lambda > 1\} \\ &= \bigcap_{\lambda>1} \lambda K\end{aligned}$$

proving (1.19). Since this set is convex, Theorem 1.9 shows $\|\cdot\|$ is a seminorm.

Similarly,

$$\begin{aligned}\{x \mid \|x\| < 1\} &= \{x \mid \inf\{\lambda \mid \lambda^{-1}x \in K\} < 1\} \\ &= \{x \mid \lambda^{-1}x \in K \text{ for some } \lambda < 1\} \\ &= \bigcup_{\lambda<1} \lambda K\end{aligned}$$ □

Gauges of balanced, absorbing, convex sets will play a key role in the theory of locally convex spaces; see Chapter 3.

Theorem 1.9 shows that there is an interplay between subadditive functions, convex functions, and homogeneous functions of degree one. There is an analog for sets.

Definition Let V be a vector space. A *cone* in V is a subset $K \subset V$ so that $x \in K$ and $\lambda \ge 0$ implies $\lambda x \in K$. K is called additive if and only if $x, y \in K$ implies $x + y \in K$.

The following analog of Theorem 1.9 is immediate:

Theorem 1.11 *Let K be a cone in V. Then K is convex if and only if K is additive.*

It is also easy to see that if K is both convex and additive and $0 \in K$, then K is a cone.

Convex cones are often easier to deal with than convex sets. For this reason, given any convex $K \subset V$, we define the set

$$K_{\text{sus}} = \{(\lambda x, \lambda) \mid x \in K,\ \lambda \ge 0\} \subset V \times \mathbb{R} \tag{1.21}$$

called the *suspension* of K. It is easy to see that K_{sus} is a convex cone if and only if K is convex.

Example 1.12 (Orlicz Spaces and Minkowski's Inequality) Let $(M, d\mu)$ be a measure space with $\mu(M) \equiv 1$. ($\mu(M)$ finite is easily handled, μ σ-finite is harder, but many of the results in the next chapter – suitably modified – hold.) Let F be a convex function on $[0, \infty)$ with $F(0) = 0$ and $F(y) > 0$ for all $y > 0$. We suppose that $\lim_{y\downarrow 0} F(y) = 0$. We will use the fact proven below (see Theorem 1.19) that

Cambridge University Press
978-1-107-00731-4 - Convexity: An Analytic Viewpoint
Barry Simon
Excerpt
More information

F is continuous. For any measurable function f on M, define

$$Q_F(f) = \int F(|f(x)|)\, d\mu(x) \tag{1.22}$$

where Q_F may be $+\infty$. Then, because F is convex, $Q_F(\cdot)$, where finite, is convex and thus,

$$K = \{f \mid Q_F(f) \leq 1\}$$

is a convex set which clearly obeys $-K = K$ since $Q_F(-f) = Q_F(f)$.

Define $\tilde{L}^{(F)}(M, d\mu)$ by

$$\begin{aligned} &\tilde{L}^{(F)}(M, d\mu) \\ &\qquad = \{f \text{ measurable from } M \text{ to } R \mid Q_F(\alpha f) < \infty \text{ for some } \alpha > 0\} \end{aligned} \tag{1.23}$$

Clearly, $\tilde{L}^{(F)}$ is closed under scalar multiplication. Moreover, if $\gamma = (\alpha^{-1} + \beta^{-1})^{-1}$, we have

$$Q_F(\gamma(f+g)) \leq \gamma\alpha^{-1} Q_F(\alpha f) + \gamma\beta^{-1} Q_F(\beta g) \tag{1.24}$$

since $\gamma(f+g) = \gamma\alpha^{-1}(\alpha f) + \gamma\beta^{-1}(\beta g)$, F is convex, and $\gamma\alpha^{-1} + \gamma\beta^{-1} = 1$. By (1.24), $\tilde{L}^{(F)}$ is closed under sums, so $\tilde{L}^{(F)}$ is a vector space.

Note that if $Q_F(\alpha f) < \infty$ for some α, then by the monotone convergence theorem ($F(x)$ is monotone on $[0, \infty)$ by hypothesis),

$$\lim_{\alpha \downarrow 0} Q_F(\alpha f) = 0 \tag{1.25}$$

Moreover, if f is not a.e. 0,

$$\lim_{\alpha \to \infty} Q_F(\alpha f) = \infty \tag{1.26}$$

(It may happen $Q_F(\alpha F) = \infty$ for some $\alpha < \infty$.) By (1.25), hypothesis (i) of Corollary 1.10 holds.

Thus, Corollary 1.10 lets us construct a seminorm

$$\|f\|_F = \inf\{\lambda > 0 \mid Q_F(\lambda^{-1} f) \leq 1\} \tag{1.27}$$

called the *Luxemburg norm* associated to F. By (1.23), we get a norm by taking equivalence classes of functions equal a.e. We call this space the Orlicz space associated to F and denote it by $L^{(F)}(M, d\mu)$.

If $F(x) = |x|^p$, then

$$Q_F(f) = \int |f(x)|^p\, d\mu(x)$$

and $Q_F(\lambda^{-1} f) = \lambda^{-p} Q_F(f)$. Therefore, $Q_F(\lambda^{-1} f) \leq 1$ if and only if